Real-Time Howling Detection for Hands-Free Video Conferencing System

Mi Suk Lee and Do Young Kim

Future Internet Research Department

ETRI, Daejeon, Korea

{lms, dyk}@etri.re.kr

Abstract*:*

This paper presents howling detection method in two-way hands-free communication system. Though hands-free communication is convenient for multi-user system such as audio/video conference, it is very common for the open microphones and loudspeakers produce acoustic feedback in a closed loop, which results in howling. This obviously prevents any useful conversation between participants.

In this paper we propose a real time howling detection method based on the long-term average spectral power and howling information of the previous frame. Performance tests show that the proposed algorithm provides quick and stable howling detection results.

Keywords-component; *howling detecion; howling suppression; hands-free video conferencing*

I. INTRODUCTION

As high speed internet connectivity has become more easily available, video conferencing is being used more and more everyday life for business and educational purpose, not just for the personal use. Recently, many educational institutions including universities are interested in distance education by using the video conferencing technology for students who are separated by time and distance. In the fields of business, various types of video conferencing system has been adopted, because it enable individuals in distance locations to participate in meetings with time and money saving.

In video conferencing system, hands-free communication is more convenient for users. But it is very common to produce acoustic feedback in a closed loop, resulting in howling. This obviously prevents any useful conversation between participants. A typical closed loop path (dashed line) can exist when both parties are using hands-free video conferencing system is illustrated in figure 1. This type of closed loop which includes microphone, network, and loudspeaker is most common in real time hands-free communication system. If the loop gain is greater than unity at one or more frequencies the system becomes unstable and produces oscillations.

If the oscillation signal has the frequency in the audible range of human hearing, it is called howling or squealing. It not only disturbs normal communications, but also damages power amplifier for overload.

In two-way hands-free communication, acoustic feedbacks from loudspeaker into the microphone are traditionally cancelled using acoustic echo canceller (AEC). When it has reached enough convergence, it provides protection against howling because it reduces acoustic feedback. Until then, or when the acoustic path is changed or high loudness is required, the system exposed to howling because of improper function of AEC[1,2]. One of This research was funded by the MSIP (Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2013. The most popular techniques for supplementing the AEC to prevent howling in hands-free communication systems is notch filter based howling suppression (NHS) methods.

NHS methods consist of howling detection and howling suppression by using a notch filter. The design of notch filters for howling suppression is based on well-established filter design techniques. Hence, it is critical to detect howling signal rapidly and exactly in NHS methods. In this paper we propose a howling detection method based on the howling information of the previous frame and long-term average spectral power for real time twoway hands-free video conferencing system.

This paper organized as follows: Section 2 briefly reviews a notch filter based howling suppression method. In section 3, we discuss characteristic of howling signal recorded in our two-way hands-free video conferencing system. Section 4 introduces a proposed howling detection method for NHS and Section 5 presents performance test results.

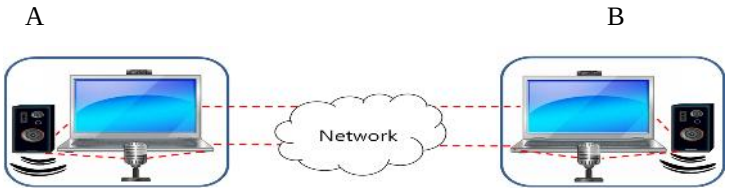

Figure1. Acoustic feedback in packet based video conferencing system

II. NOTCH FILTER BASED HOWLING SUPPRESSION

Notch filter based howling suppression (NHS) is one of the most popular methods for acoustic feedback control in public address and hands-free communication systems[3]. The NHS methods consist of howling detection and howling suppression by using a notch filter. And NHS can be divided into two categories, i.e., onestage and two-stage, depending on whether the howling detection and notch filtering are performed jointly or separately. The adaptive notch filter (ANFs) based methods are typical one-stage method.

Two-stage NHS methods are the most popular method for acoustic feedback control. The two-stage NHS in hands-free communication system can be outlined as shown in figure 2. The microphone signal is first processed by a howling detection algorithm, which forwards a set of design parameters of notch filter. When howling has been detected, a notch filter has been activated to suppress the howling signal before transmit the input signal to the remote site. In the two-stage NHS method, howling detection is crucial to get reliable howling suppression results because the notch filter design method is well defined.

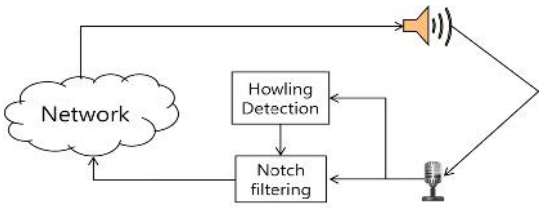

Figure 2. Two-stage notch filter based howling suppression

III. HOWLING SIGNAL IN TWO-WAY HANDS-FREE VIDEO CONFERENCING ENVIRONMENT

In order to analyze the howling signal in two-way hands-free video conferencing environment, we set up an experimental environment as shown in figure 1. The hands-free video conferencing system consists of personal computer, display, microphone and loudspeaker. The characteristic of the microphone and loudspeaker is as follow:

- Loudspeaker: BOSE Companion 2 multimedia speaker
- Microphone: ETM-003 of Edutige

- Condenser / Omnidirectional / Boundary microphone
- Sensitivity: -23dB
- Frequency response: 50~18 kHz

Figure 3 shows a block diagram for audio signal processing in our video conferencing system. The microphone signal of site A is processed by AEC and encoded with G.711.1 wideband encoder [4] and transmitted to remote site B after RTP packing. In site B, the received packet is unpacked and decoded by G.711.1 decoder. The decoded signal is played out through a loudspeaker. Since the loudspeaker and microphone are located in same room, the speaker output is captured by a microphone again and transmitted to site A.

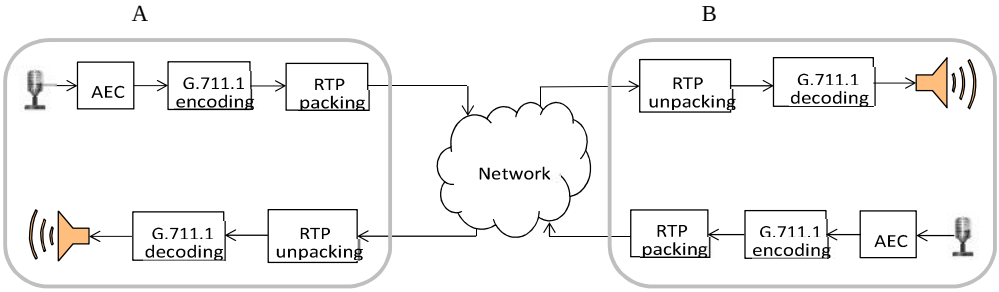

Figure 3. Audio signal flow in two-way hands-free video conferencing

Figure 4 and 5 present time domain signal and frequency spectrum of howling signal recorded at 16 kHz sampling rate without AEC block both in site A and B, respectively. The frequency range of each spectrum in those figures is limited to 1,000 Hz. From the figures, we can see the amplitude of howling signal is globally increase with time but has some fluctuation.

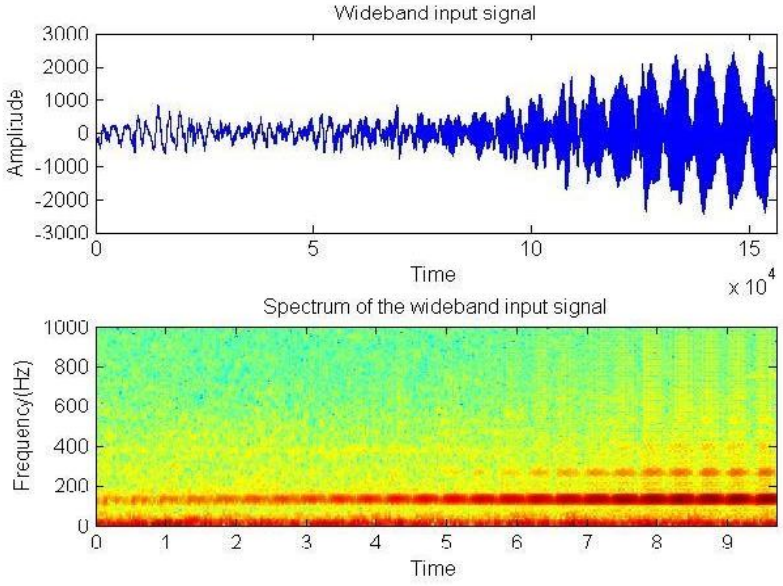

Figure 4. Howling signal recorded in site A

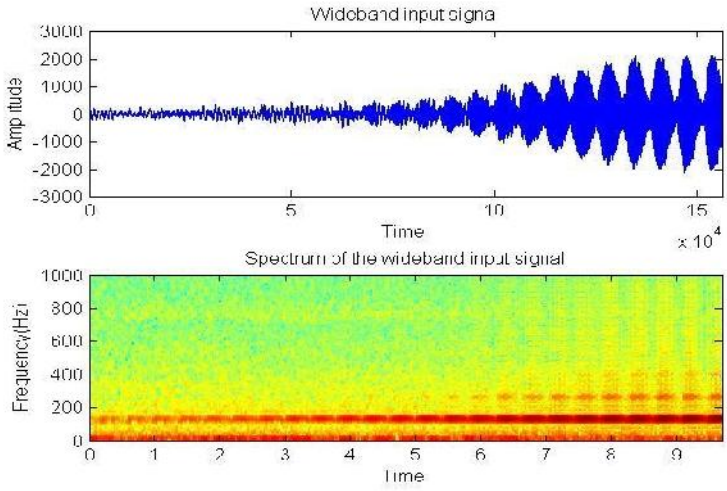

Figure 5. Howling signal recorded in site B

IV. PROPOSED HOWLING DETECTION METHOD

Figure 6 presents the high-level block diagram of the proposed howling detection method. Basically, it is operated in a 10 msec frame based manner. The microphone input signal is stored in a buffer and this time domain signal is transformed into a frequency domain signal by a FFT analysis. FFT is run once every 10 msec and the FFT window size is 16 msec (corresponding to 256 samples at 16 kHz sampling). After howling candidate selection and detection, the related parameters are update for the next frame processing.

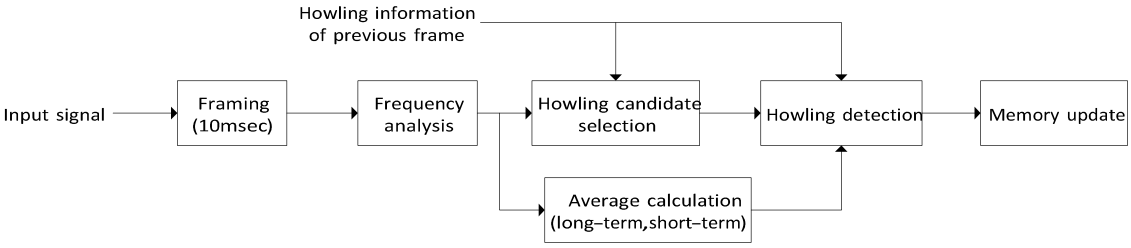

Figure 6. High level block diagram of the proposed howling detection

A. Howling Candidate Selection

The magnitude of the howling signal is globally increased over time, but it can be locally fluctuated by the delay caused in communication environment. In this case, it is difficult to detect howling signal consistently in some segment which have relatively decreased magnitude. In order to solve this problem, we use the howling information of the previous frame when selecting howling candidate.

To select a howling candidate, first, N frequencies which have highest spectral power are chosen by using peak picking. We select a frequency which has maximum spectral power from N as a howling candidate of the current frame. However, if the selected howling candidate is different from the howling frequency of the previous frame, the howling candidate can be replaced with the howling frequency of the previous frame according to the following procedure.

First, it is checked that there is a same frequency with the howling frequency of the previous frame in N-1 frequencies except howling candidate. Second, if one of them is same with the howling frequency of the previous

frame, then compare the spectral power of the current howling candidate and the previous howling frequency. The howling candidate of the current frame are replaced with the howling frequency of the previous frame if the spectral power at previous howling frequency is greater than the weighted spectral power of current howling candidate. Here, the weighing factor is set to a value less than 1.

B. Howling Detection

The howling detector judges that the howling candidate is a true howling signal. We use not only the peak to short-term (frame) average spectral power ratio but also long-term average spectral power ration of the howling candidate of current frame and howling frequency of the previous frame.

We detect the howling candidate as a true howling signal in two cases. First, it is decided the howling frequency candidate is true howling signal when both ratio of the spectral power of the howling candidate to a short-term average spectral power and the long-term to short-term average spectral power ratio of the howling candidate is greater than the predetermined threshold values. Second, it is detected the howling candidate is a true howling signal when the previous frame has howling signal and the long-term average spectral power of the howling candidate is greater than the weighted long-term average spectral power of the howling signal of previous frame. The weighing factor is set to a value less than 1.

V. EXPERIMENTAL RESULTS

In order to test the performance of our proposed method, we recorded two types of signal without AEC operation at 16 kHz sampling rate. One is a pure howling signal without speech activity and second is a howling signal with a speech activity from one of the sites. Figure 7 shows a test result for a pure howling signal in time domain and frequency domain with a limited frequency range.

A microphone input signal (blue wave) and howling suppressed signal (green wave) by using a 2nd order IIR notch filter is shown in figure 7. And the howling detection results also depicted with black line overlapped with the waveform and spectrum. Figure 7 shows that the proposed method is able to detect the howling signal very quickly and consistently even though the magnitude of howling signal is fluctuated.

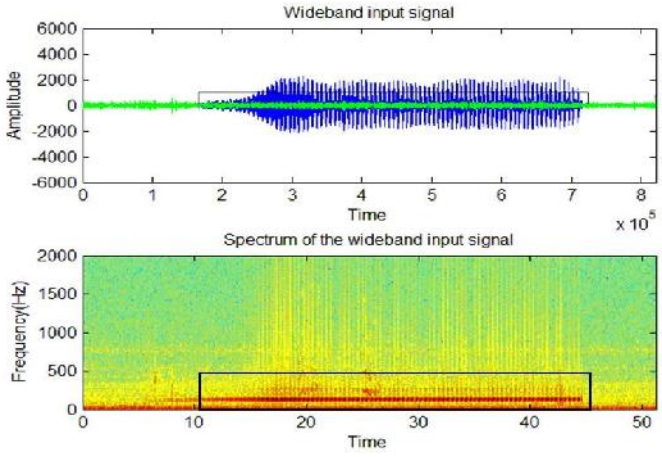

Figure 7. Experimental results for a pure howling signal.

Figure 8 shows a performance test results for a howling signal mixed with speech signal. The time domain waveform of input signal(blue wave) and howling suppressed signal(green wave) are presented with howling

detection results(black line) in first block of figure 8. Also, the spectrum of the microphone input signal and howling suppressed signal is illustrated with limited frequency range. It indicates that the proposed howling detection is providing stable performance when the speech signal is presented.

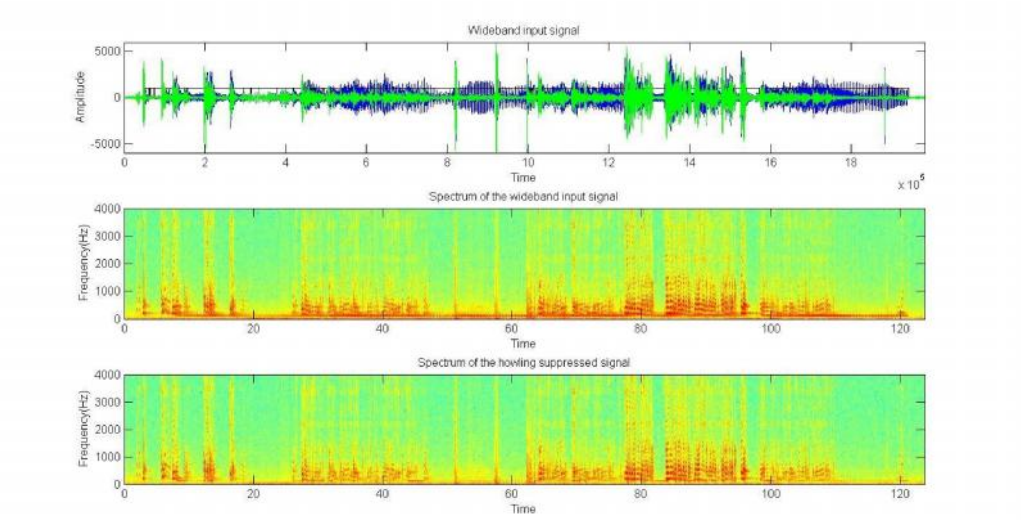

Figure 8. Experimental results for a howling signal mixed with speech signal.

VI. CONCLUSION

In this paper, we propose a howling detection method based on the howling information of previous frame and long-term average spectral power. The performance of the proposed method is tested with the sample data recorded in real-time two-way hands-free video conferencing environment. The test results shows that the proposed method quickly detect the howling frequency and also gives stable results for the howling signal mixed with speech.

REFERENCES

[1] Senthil Kumar Mani and Sowmya Mannava, "Robust and High quality Howling Suppression for Real Time Hands-Free Communication systems", IEEE 2011

[2] Franck Beaucoup, " A Novel Loop Stabilisation Technique for Full-Duplex Speakerphones", IEEE CCECE/CCGEI, May, 2006, pp.771-774.

[3] T.van Waterschoot and M.Moonen, "comparative spectral analysis of howling detection criteriia in notch filter based howling suppression". AES 126th conv., Munich, Germany, May, 2009

[4] ITU-T Recommendation G.711.1, "Wideband embedded extension for ITU-T G.711 pulse code modulation ", sept. 2012.